# The relationship between historical development and potentially toxic element concentrations in urban soils

McIlwaine, R., Doherty, R., Cox, S. F., & Cave, M. (2017). The relationship between historical development and potentially toxic element concentrations in urban soils. *Environmental Pollution* , *220*, 1036-1049.
https://doi.org/10.1016/j.envpol.2016.11.040

**Published in:**
Environmental Pollution

**Document Version:**
Peer reviewed version

Title: The relationship between historical development and potentially toxic element concentrations in urban soils

Authors: Rebekka McIlwaine[1]

Rory Doherty[1]

Siobhan F. Cox[1]

Mark Cave[2]

[1] Civil Engineering Research Centre
School of Planning, Architecture and Civil Engineering
Queen's University Belfast
Belfast BT7 1NN
Northern Ireland, UK

[2] British Geological Survey
Keyworth
Nottingham NG12 5GG
England, UK

Abstract

Increasing urbanisation has a direct impact on soil quality, resulting in elevated concentrations of potentially toxic elements (PTEs) in soils. This research aims to assess if soil PTE concentrations can be used as an 'urbanisation tracer' by investigating geogenic and anthropogenic source contributions and controls, and considering PTE enrichment across historical urban development zones. The UK cities of Belfast and Sheffield are chosen as study areas, where available shallow and deep concentrations of PTEs in soil are compared to identify geogenic and anthropogenic contributions to PTEs. Cluster analysis and principal component analysis are used to elucidate the main controls over PTE concentrations. Pollution indices indicate that different periods of historical development are linked to enrichment of different PTEs. Urban subdomains are identified and background values calculated using various methodologies and compared to generic site assessment criteria. Exceedances for a number of the PTEs considered suggest a potential human health risk could be posed across subdomains of

both Belfast and Sheffield. This research suggests that airborne diffuse contamination from often historical sources such as traffic, domestic combustion and industrial processes contribute greatly to soil contamination within urban environments. The relationship between historical development and differing PTEs is a novel finding, suggesting that PTEs have the potential for use as 'urbanisation tracers'. The investigative methodology employed has potential applications for decision makers, urban planners, regulators and developers of urban areas.

Capsule

Potentially toxic elements have the potential for use as 'urbanisation tracers' due to their association with different historical anthropogenic sources.

# 1 Introduction

Globally, more people now live in urban areas than in rural areas; in 2014 54% of the world's population lived in urban areas. This has rapidly increased from just 30% in 1950 and is projected to reach 66% by 2050 (UN 2014). The demand put on these geographically limited urban environments will intensify as population density increases, with a direct impact on soil quality.

Sources of potentially toxic elements (PTEs) in urban areas are often both geogenic and anthropogenic (Argyraki & Kelepertzis 2014; Rodrigues et al. 2009), with both point and diffuse anthropogenic sources common (Marchant et al. 2011). Typical anthropogenic sources such as industry, traffic (leaded fuel, brake pads and tire wear (Argyraki & Kelepertzis 2014; Dao et al. 2014)) and waste disposal are known to contribute to PTE concentrations in soil (Ajmone-Marsan & Biasioli 2010). Domestic outputs in urban environments, in the form of fuel burning and waste, can also be large contributors to soil PTE concentrations (Biasioli et al. 2006; Glennon et al. 2014).

As urban areas continue to grow, a factor that must be considered is how human health can be affected by PTEs in soil. Can development be appropriately and sustainably managed considering previous land uses and soil PTE concentrations? The economic importance of urban soils must be balanced with ensuring potentially contaminated urban sites are safe for redevelopment.

Numerous urban geochemical investigations have been completed across the world (Biasioli et al. 2006; Argyraki & Kelepertzis 2014; Glennon et al. 2014; Kelepertzis & Argyraki 2015; Golden et al. 2015;

Johnson et al. 2011; Thorton 2009; Mielke 1999) and it is by building upon this library of research that we can fully understand how urban PTE sources vary geographically. The study areas used in this research have diverse bedrock and rich industrial histories, making them the ideal locations for investigating combined geogenic and anthropogenic contributions to soil PTE concentrations.

This research aims to understand if PTE concentrations in soil can be used as a tracer for urbanisation by; (1) investigating geogenic and anthropogenic contributions to PTE concentrations in soil, (2) identifying groups of PTEs controlled by similar sources, (3) understanding how historical city development may have influenced soil quality by considering PTE enrichment across city development zones and (4) calculating typical threshold values for the anthropogenic PTEs from similar sources. A novel investigative methodology will be employed utilising depth ratios, a range of multivariate statistical techniques and pollution indices. The objective is to generate a methodology for use in other urban areas, for a range of potential pollutants, to inform on city areas most likely to be contaminated.

Eleven PTEs are considered; arsenic (As), cobalt (Co), chromium (Cr), copper (Cu), molybdenum (Mo), nickel (Ni), lead (Pb), antimony (Sb), tin (Sn), vanadium (V) and zinc (Zn). These elements are expected to be related to different geogenic and anthropogenic sources within the study areas; in particular they are likely to represent contamination from a variety of historical industrial processes. Elements such as As, Cr, Cu, Ni, Pb and Zn commonly feature in urban geochemical studies due to their anticipated anthropogenic sources (Johnson & Ander 2008) while Carrero et al. (2013) demonstrate a relationship between a variety of the chosen PTEs, including Mo, Sb and Sn, and soils heavily impacted by traffic. Previous research in one of the study areas (McIlwaine et al. 2014; McIlwaine et al. 2015; Cox et al. 2013; Barsby et al. 2012; Palmer et al. 2015) has demonstrated concentrations of various PTEs to be controlled by geogenic sources in the form of bedrock geology.

# 2 Methodology

## 2.1 Study areas

Two cities within the United Kingdom have been chosen as study areas for this research. Belfast, Northern Ireland's capital and largest city has been compared with Sheffield, a city in South Yorkshire, England. These cities were selected due to their similar industrial heritage in heavy engineering (although Sheffield's is slightly greater and more varied than Belfast) and their similar geology.

The Corine land cover data (European Environment Agency 2012) was utilised to define both the Belfast and Sheffield study area boundaries (Figure 1). This data is split into 44 different land uses based on the interpretation of satellite images. The majority of land uses were simplified into either urban or rural; areas of urban fabric and industrial units were obviously urban while areas of pasture and forest were clearly rural. Where any inconclusive land uses were identified, the surrounding land use was utilised as an indicator of land use type on a site by site basis. Within Sheffield, the Corine boundary was slightly reduced due to the spatial extent of the geochemical data available.

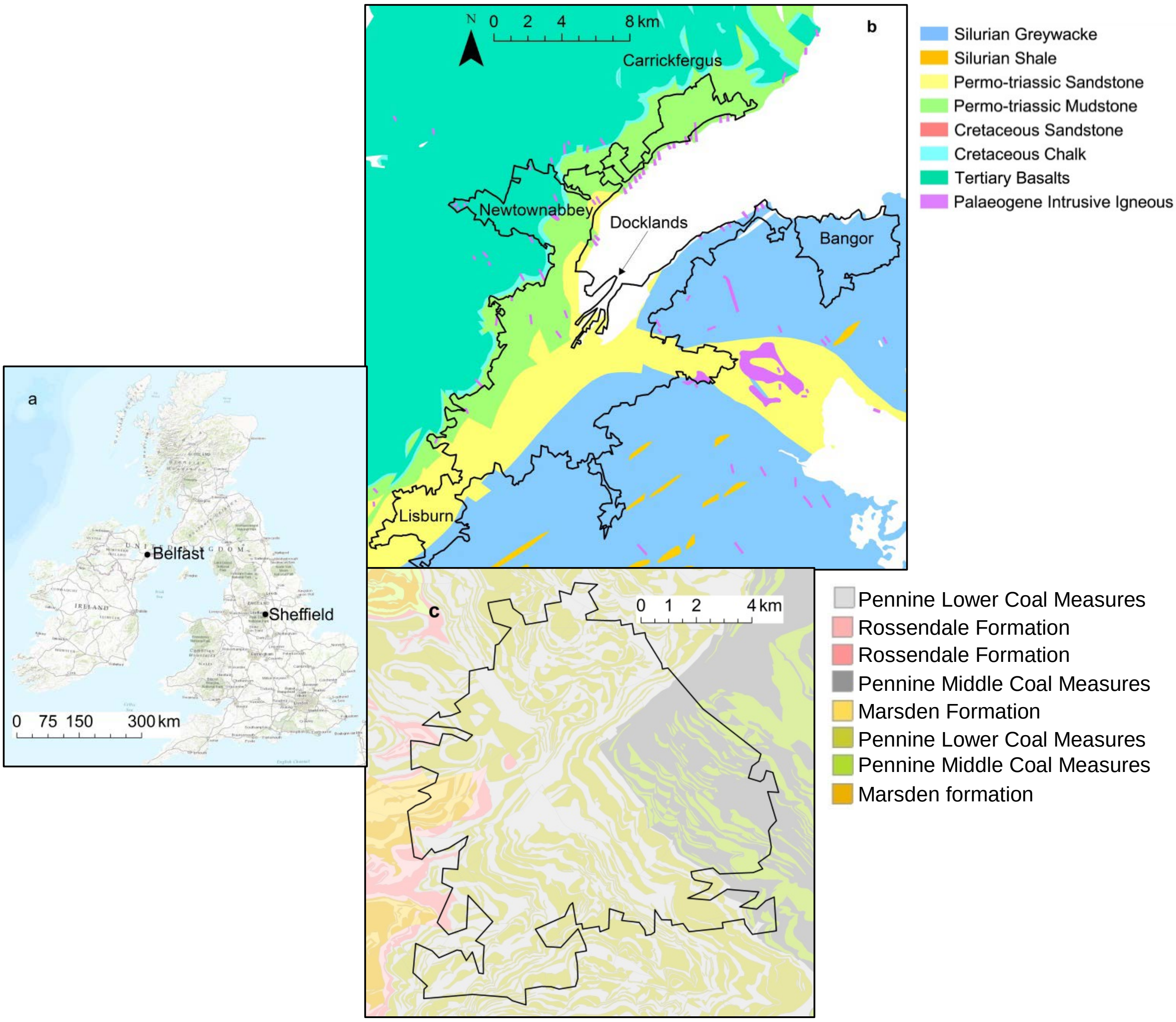


*Figure 1 Maps showing a) location of Belfast and Sheffield within the UK, b) simplified bedrock geology in Belfast (bedrock geology derived from data provided by Geological Survey Northern Ireland (GSNI) (Crown Copyright)) and c) bedrock geology in Sheffield (taken from BGS GeoIndex)*

### 2.1.1 Belfast

A simplified representation of the highly diverse bedrock geology in Belfast, from GSNI's 1:250000 map, is produced in Figure 1. Silurian greywacke and Silurian shale are the two oldest rock formations, forming part of the Southern Uplands-Down-Longford Terrane, followed by the Permo-triassic sandstones and mudstones. This is covered in the west of the city by Cretaceous sandstone and chalk and finally by the most recent Tertiary basalts which run along the north west boundary of city (Mitchell 2004). A number of Palaeogene intrusions occur within the study area. Geology has previously been identified as a control over element concentrations in Northern Irish soil (McIlwaine et al. 2014), with areas of basalt and sandstone identified as containing elevated concentrations of different PTEs. This study area is therefore of interest due to the expected geogenic controls within an urban environment.

Superficial geology within Belfast (Supplementary Information 5) has been reproduced from a GSNI map showing the geology of Belfast and District (Bazley et al. 1984). It is found in the form of till, glacial sands and gravels, and alluvium within the vicinity of the River Lagan.

Historical maps of Belfast (Land and Property Services 1858; Land and Property Services 1901; Land and Property Services 1919) have been used to produce historical study area boundaries for 1858, 1901 and 1919-1939. Historical development of surrounding towns that are now incorporated in the greater Belfast area (Carrickfergus and Bangor on Figure 1) has not been included.

Belfast is historically recognised for both linen production and ship-building; the early 18th century saw the introduction of the linen industry involving bleaching, weaving and spinning processes while ship-building was introduced later in the 18th century (Beckett & Glasscock 1967; Crawford 1986). The city was an important manufacturing centre during the industrial revolution with other common businesses including rope works, bleachers, glass manufacturers, tobacco factories and distilleries (Royle 2007). Present-day Belfast is much more reliant on service provision related occupations (82% in Northern Ireland in the 2013 Census of Employment (NISRA 2014)) than the historical industrial and manufacturing employment.

### 2.1.2 Sheffield

Sheffield is underlain by Carboniferous deposits of Westphalian and Namurian age (Freestone et al. 2004). The rocks are highly faulted and folded creating many discontinuous outcrops (Figure 1). The Middle Coal Measures Formation outcrops to the east of the city centre; it is Westphalian in age and composed of sandstone. The Lower Coal Measures Formation, also composed of sandstone and Westphalian in age, underlies most of the city centre. The Rossendale and Marsden Formations are present to the west of the study area comprising of mudstone and siltstones.

Quaternary deposits cover approximately 10% of Sheffield (Supplementary Information 5); silt alluvium is located in the river valleys around and in the city of Sheffield while some river terrace deposits in the form of sand and gravel also infill these valleys (Freestone et al. 2004).

Historical maps (OS Six Inch England and Wales 1851; Bartholomew's "Half Inch Maps" of England and Wales 1904; Bartholomew's Revised "Half Inch Maps" 1920; 1:25000 maps of Great Britain 1953) of

Sheffield have also been used to produce historical study area boundaries for 1850-1851, 1904, 1920 and 1938-1951.

Sheffield is located in South Yorkshire in England and is believed to have been originally founded in the 8th century. Coal has played an important role in the city's history, being mined and burnt for "space heating and industrial purposes in Sheffield since Roman times" (Rawlins et al. 2005). By the 1750s, more than 150 firms were dedicated to steel manufacture within the city. High quality cutlery, an export for which Sheffield is recognised, has been produced in the city since that time (Gilbertson et al. 1997). Industrial expansion continued until the late 1960s, when British Steel opened their Tinsley Park Works in the north east of the city. The steel and cutlery industry in Sheffield began to decline in the late 1970s and 1980s when cheaper alternatives were being produced in other areas of the world. This also affected the coal used to fuel industry in Sheffield, with coal use declining dramatically from the mid-1980s (Gilbertson et al. 1997).

## 2.2 Geochemical data

The Tellus project, managed by GSNI, was completed across Northern Ireland between 2004 and 2007, comprising both geophysical and geochemical surveys. In total, 781 Tellus soil sample locations (4 samples per $km^2$) fall within the defined Belfast boundary (<200 $km^2$). Geochemical measurements are available via a variety of different analytical techniques. Total concentrations determined by X-ray fluorescence (XRF) were available for the shallow (5-20 cm) soil samples. Aqua regia extractable data are available for the shallow and deep (35-50 cm) samples, with the PTEs investigated within this study analysed by a mixture of Inductively Coupled Plasma (ICP) Optical Emission Spectrometry (OES) and ICP Mass Spectrometry (MS).

Similar sampling density and analytical techniques were employed during G-BASE sampling of Sheffield. However, this G-BASE data was solely analysed by XRF and at this stage only the shallow (5-20cm) soils have been analysed. Some 495 G-BASE soil sample locations fall within the defined Sheffield boundary. More detailed information on the Tellus datasets including quality assurance and quality control procedures can be found in Green et al. (2010); Knights (2007) and Smyth (2007) and information on the Geochemical Baseline Survey of the Environment (G-BASE) protocols followed are provided by Johnson (2005).

## 2.3 Depth comparison

Comparing shallow (5-20cm) and deep (35-50cm) PTE concentrations gives a greater understanding of anthropogenic and geogenic inputs to PTE concentrations in soil. Generally, if the shallow concentrations are more elevated this suggests an anthropogenic control over the PTE (Chiprés et al. 2009) while a geogenic control will result in elevated concentrations in the deep samples (Galán et al. 2008). Although this assumption will generally hold true, it is important to note that urban soil is often replaced or altered due to development, with potential for contaminated soils to be reworked or placed at depth. Even with relatively undisturbed soils care should be taken as the presence of organic matter may cause differences between horizons (Reimann & Garrett 2005).

Ratio boxplots based on the shallow concentration of the PTE divided by its deep concentration have been constructed to provide an indication of the general inputs to different PTEs in Belfast. The data based on the aqua regia extraction, followed by an ICP finish was used for this comparison as it is available at both depths therefore providing a valid comparison. This comparison could not be completed for Sheffield where only shallow data is available. The docks area of Belfast (Figure 1) is developed on reclaimed land so, within that small area (approximately 6.5 km$^2$), the assumptions regarding depth are not likely to hold true, however, there is no particular evidence to suggest this occurs elsewhere within Belfast. The scale of the data used to complete this comparison and the size of the study area suggests that this depth comparison will inform on controlling sources.

## 2.4 Source identification and relationships between PTEs

Multivariate techniques in the form of cluster analysis and principal component analysis (PCA) have been used to determine underlying controls over the PTE dataset (Andersson et al. 2010; Candeias et al. 2011; Argyraki & Kelepertzis 2014).

Shallow XRF data was used for this analysis; previous research by McIlwaine et al (2015) suggests elemental form may affect the concentrations determined by ICP-OES/ICP-MS following an aqua regia digestion. This is of particular importance within urban environments as the source of the PTE is likely to control its form. Shallow data are of more relevance for this section of the research due to interest in the anthropogenic controls over the PTEs, and also because shallow/surface soils drive risks to humans through inhalation of dust, ingestion of soil and dermal contact (Cole & Jeffries 2009; CL:AIRE 2014a; Nathanail et al. 2015).

Geochemical data is compositional in nature, meaning that all the values are relative to each other i.e. all the elements analysed in a sample sum to a constant value. Therefore, all total element concentrations depend on the concentrations of the other elements in that sample meaning that they should be 'opened' prior to multivariate statistical analysis (Pawlowsky-Glahn & Egozcue 2006; Reimann et al. 2012; Aitchison 1982). The centred log-ratio (clr) transformation was found to be an appropriate manner for 'opening' geochemical data prior to multivariate data analysis within this research as it allows retention of the relationship with the original variables of the dataset. Prior to completion of the cluster analysis and PCA the data was also scaled to unit variance to ensure differences in scale would not control the outputs.

Cluster analysis has been used as an exploratory data analysis method with the aim of splitting the data under consideration into a number of groups which are similar in their characteristics or behaviour (Reimann et al. 2008). The commonly utilised Ward's minimum variance method (Astel et al. 2007; Frentiu et al. 2013; Ward 1963) was used to form groups of subsets based on their similarity as defined by specified characteristics and the Euclidean distance.

The PCA plots geochemical data in multivariate space, searching for the direction that contains maximum variability. The resulting loadings describe the relationship between the original variables and the Principal Components (PCs), while the scores describe the relationship between each of the observations and the PCs (Reimann et al. 2008). In order to understand the spatial distribution of the PCs, the score for each observation has been plotted, forming a map of how the PC is distributed. This allows a comparison between controlling variables and geographical distribution, providing a full interpretation of the PCA (Reimann et al. 2008). This interpolated map was produced using inverse distance weighting (output cell size of 200m, power of 2 and a search radius of 500m) which is recommended for use within spatially dense networks (Dirks et al. 1998).

## 2.5 Urban growth

Pollution indices (PIs) have been utilised to gain an understanding of how the PTE concentrations within the Belfast and Sheffield development zones (based on the historical boundaries discussed in Section 2.1) show various levels of enrichment. These PIs have been calculated using;

$$PI = U_c/R_c$$

where $U_c$ is the median element concentration within the development zone under consideration and $R_c$ is the median rural element concentration. A study by Biasioli et al. (2006) used PIs in a similar manner to estimate the enrichment of a city with certain PTEs. In order to determine the influence of different urban environments over PTE concentrations through anthropogenic sources the most effective comparison is with rural environments. Previous research has suggested that there was no obvious agricultural input of anthropogenic PTEs to the rural environment within Northern Ireland (McIlwaine et al. 2014; McIlwaine et al. 2015). The median rural concentration has been used as the median is more robust to outlying values than the more commonly used mean (Zhang et al. 2007).

The median rural element concentrations were easy to calculate for the Belfast study area, as the Tellus rural data was available. The median rural element concentrations for the Sheffield study area were taken from Freestone et al. (2004), which provided the 'median concentrations in regional surface soil samples overlying Carboniferous Coal Measures, Humber-Trent atlas areas'.

## 2.6 Background or typical threshold values

A variety of methods for calculating background concentrations were utilised, and the results contrasted and compared to gain an understanding of their strengths. The domains (areas where a readily identifiable factor can be shown to control the concentration of the PTE) were based on the findings of the PCA and historical analysis, and as such were defined using the historical development zones identified. As previously suggested for use at a regional scale in Northern Ireland, the typical threshold value (TTV) (McIlwaine et al. 2014) methodology which utilises the Finnish upper limit of geochemical baseline variation (ULBL) (Jarva et al. 2010) to define background concentrations was employed. The ULBLs are based on the upper limit of the upper whisker line of the box and whisker plots, which can be calculated using:

$$\text{ULBL} = P_{75} + 1.5 \times (P_{75} - P_{25})$$

$P_{75}$ and $P_{25}$ are the 75th and 25th percentiles of the element concentrations respectively (Jarva et al. 2010). Logarithmic transformed data were not used to plot the box and whisker plots, as the untransformed data led to the highest amount of outliers and therefore gives a more conservative value. This is a very straightforward method for calculating background values, solely requiring the calculation of the 75th and 25th percentiles.

In addition, the Normal Background Concentration (NBC) method employed for use alongside Part 2A of the Environmental Protection Act in England and Wales was used. Within the NBC methodology, it is recommended that the domains are based on at least 30 values (Cave et al. 2012). The NBC is calculated for each domain using a statistical methodology given in (Cave et al. 2012). The NBC is then taken to be the upper 95% confidence limit of the 95th percentile. The project outputs included R code scripts which can be used to determine NBCs, as discussed by Johnson et al. (2012). These R scripts have been utilised by the author to calculate NBCs within this research.

The median + 2MAD (median absolute deviation) (Reimann et al. 2005) method was also utilised. The median + 2MAD, boxplot upper whisker and English NBC methods were compared as methods by which to calculate background concentrations in urban environments in a study by Rothwell & Cooke (2015). The lack of systematically collected geochemical data (no G-BASE data was available) meant that a different approach had to be taken within this study in Gateshead; site investigation data collected during the planning process was instead used. The local authority determined that the median + 2MAD method provided the preferred NBCs within this study as it consistently gave the most conservative values i.e. the lowest NBC (Rothwell & Cooke 2015).

# 3 Results and discussion

## 3.1 Depth ratio

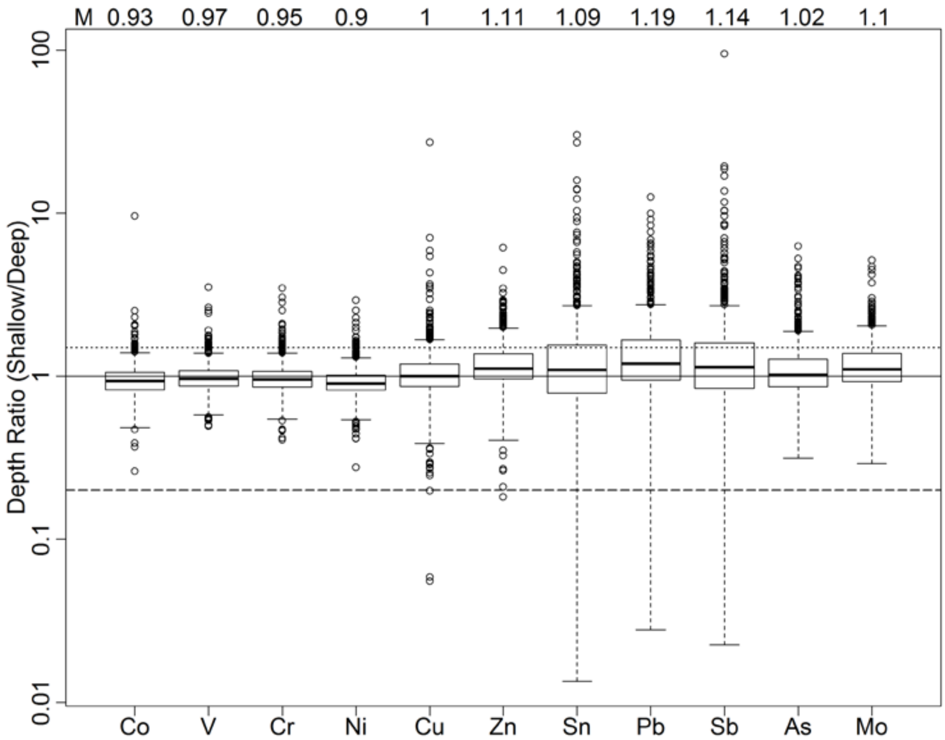


*Figure 2 Boxplots of the shallow/deep PTE concentrations (depth ratio) using ICP following an aqua regia digestion data for Belfast (solid black line shows where the shallow and deep concentrations are equal, dashed line shows where the depth ratio is equal to 0.2, dotted line shows where the depth ratio is equal to 1.5 and the M values represent the median depth ratio)*

From the Belfast depth comparison, similar boxplot characteristics are shown for Co, V, Cr and Ni; the majority of the boxplot falls below a ratio of one, representing more elevated concentrations in deep soils. A number of lower outliers are identified, related to elevated concentrations at depth, while the upper outliers are likely to be related to anthropogenic sources of these PTEs in shallow soils. This pattern suggests a predominantly geogenic control over these PTEs within Belfast. The upper outliers may be a useful way of spatially assessing anthropogenic point sources of these PTEs in the study area. These outliers show a great deal of consistency across the Co, V, Cr and Ni distributions; the depth ratio falls above 1.5 (dotted line on Figure 2) for at least 3 of the PTEs at 18 sample locations and for all 4 PTEs at 8 sample locations.

Geogenic controls are also exerted over Cu and Zn, however the overall pattern for these PTEs is a little different. A larger variance of the ratio is obvious, with a wider dispersion of the boxplot's whiskers. A number of lower outliers may represent the geogenic influence over these PTEs, however the increased amount of upper outliers suggest a more substantial anthropogenic contribution to Cu and Zn concentrations. However, it should be noted that copper can complex with organic matter in soils (Karlsson et al. 2006) which may also explain the degree of variance of the ratio observed. Further work is needed to define the effects of organic matter on such metal distribution in soils.

The remainder of the PTEs appear to be controlled by anthropogenic processes as the medians of the depth ratio are generally higher than 1 and only upper outliers are present. They can be split into two groups; As and Mo behave in a similar manner as both show a relatively small variance with only upper outliers. Although Pb, Sb and Sn also only have upper outliers they demonstrate a much larger variance. This is related to a small number of samples where the deep concentration is much greater than the shallow concentration, for example the number of samples where the depth ratio is below 0.2 (dashed line on Figure 2) is 7, 6 and 15 for Pb, Sb and Sn respectively. Most of these samples are dispersed across the study area and are potentially related to sites where the deeper soil has been disturbed or replaced (perhaps with waste materials) during development leading to higher concentrations of these PTEs at depth. An alternative possible explanation for this pattern is that these PTEs are more easily leached from shallow to deeper soils at these locations.

## 3.2 Cluster analysis

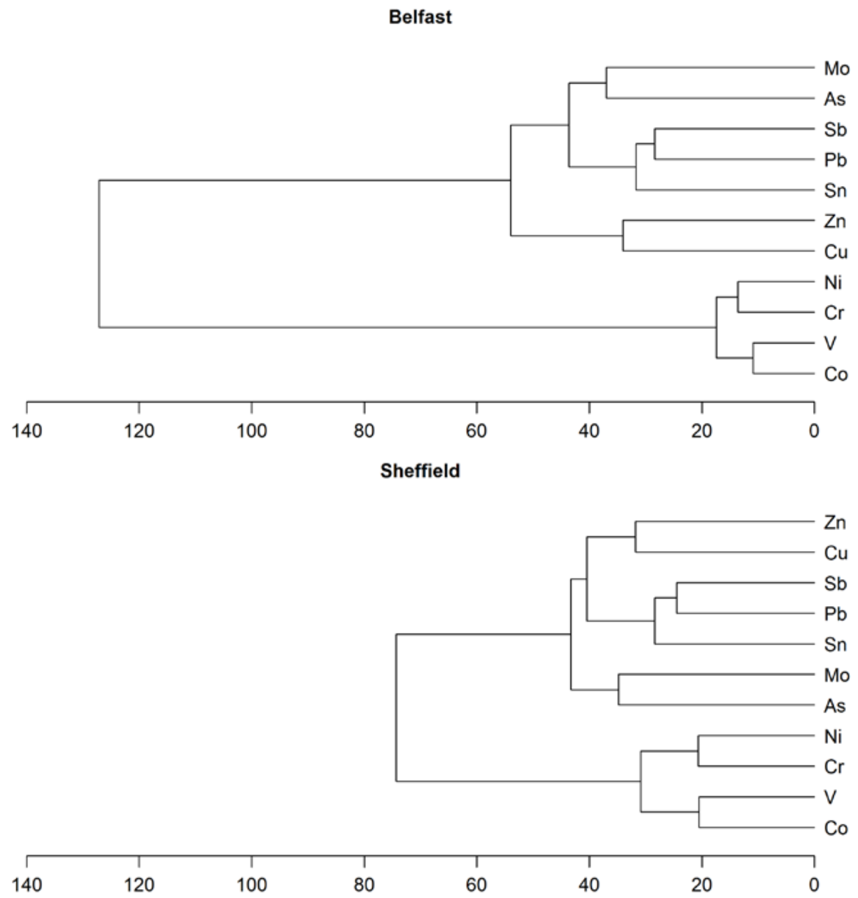


*Figure 3 Dendrograms demonstrating the cluster analysis completed for Belfast and Sheffield, with axis representing dissimilarity height between variables*

The cluster analysis (Figure 3), which considers shallow soils analysed by XRF, groups the PTEs in Belfast in a very similar pattern to that previously determined by the depth ratio analysis (Figure 2). The PTEs are first split into two main groups which can be explained by controlling geogenic (Co, V, Cr and Ni) and anthropogenic (Cu, Zn, Sn, Pb, Sb, As and Mo) factors.

Within the geogenic cluster two separate groups of Co and V, and Cr and Ni are present. Previous research has shown strong correlations between all these PTEs in a Northern Ireland context (Barsby et al. 2012), related to the stark control areas of Tertiary basalt exert over these PTEs. The total concentration maps of Co, V, Cr and Ni (Supplementary Information 1) show extremely similar spatial distributions. Although the main control over these PTEs in Belfast is the Tertiary basalts, a few anthropogenic hot spots are dispersed across Belfast. It could be these point sources, and their influence over the PTEs separately, that creates the difference between the Co and V, and Cr and Ni groups.

Three smaller groups make up the anthropogenic cluster in Belfast; firstly Cu and Zn, secondly As and Mo and finally Sn, Pb and Sb (Pb and Sb are most closely related within this cluster). As highlighted in Section 3.1, these separate groupings are related to the different source contributions for these PTEs.

Although anthropogenic controls govern these concentrations, Cu and Zn are probably grouped because they also have geogenic contributions from the Tertiary basalts. From an anthropogenic perspective, the close grouping of Cu and Zn could also be explained by their role in the production of brass (Herting et al. 2008). Arsenic and Mo are anthropogenically controlled, but again a possible geogenic influence from the Silurian greywackes resulting in elevated concentrations of As and Mo in overlying soils (Young & Donald 2013) could cause them to cluster together within the overall anthropogenic cluster. No geogenic contributions to Pb, Sb or Sn could be identified within the study area suggesting a sole anthropogenic control over these PTEs, creating their separate grouping (Figure 3).

The results of the cluster analysis for Sheffield (Figure 3) are strikingly similar to those presented for Belfast. Two main separate groupings are noted; Zn, Cu, Sb, Pb, Sn, Mo and As are grouped separately from Ni, Cr, V and Co. As demonstrated on the concentration maps (Supplementary Information 1), Ni, Cr, V and Co show elevated concentrations to the north-east of the study area. Although they are all affected by the presence of other point sources in Sheffield, the large anthropogenic source to the north-east is their most obvious characteristic, resulting in them being grouped together.

In contrast, the other PTEs all have other factors affecting their concentration distributions. This cluster analysis suggests similar sources controlling three groups of PTEs, 1) Zn and Cu, 2) Sb, Pb and Sn and 3) As and Mo. Although these elements are anthropogenically controlled it is difficult to narrow down their specific sources. As noted for Belfast, Zn and Cu could be grouped together due to their role in the production of brass. The Pb, Sb and Sn concentrations are again likely to be solely anthropogenic; the widespread nature of their elevated concentrations might suggest a controlling atmospheric deposition source.

The difference in the scale on the cluster analysis from the two cities should be noted. The overall difference between the two groups, (Ni, Cr, V and Co are grouped separately from Mo, As, Sb, Pb, Sn, Zn and Cu) is much greater for Belfast than for Sheffield. In addition, the PTEs identified as being of geogenic origin in Belfast (Ni, Cr, V and Co) have a much smaller difference in height (between the two groups and the adjacent PTEs), than is noted for these PTEs in Sheffield. This suggests stronger similarity between these PTEs in Belfast, defending the identification of controlling geogenic and anthropogenic sources of these PTEs in Belfast and Sheffield respectively.

Although the other group of PTEs (Mo, As, Sb, Pb, Sn, Zn and Cu) show similar height differences in Belfast and Sheffield between most of the groups, the overall difference between the Zn and Cu group, and the Mo, As, Sb, Pb and Sn group is much greater for Belfast than it is for Sheffield. This could be explained by the geogenic contributions to Zn and Cu in Belfast compared to the governing anthropogenic controls over Mo, As, Sb, Pb and Sn.

These results are useful for providing preliminary information on links between PTEs; in order to ascertain the underlying associations between the anthropogenic PTEs more detailed multivariate analysis is required, along with an understanding of how their sources vary spatially.

## 3.3 Principal component analysis

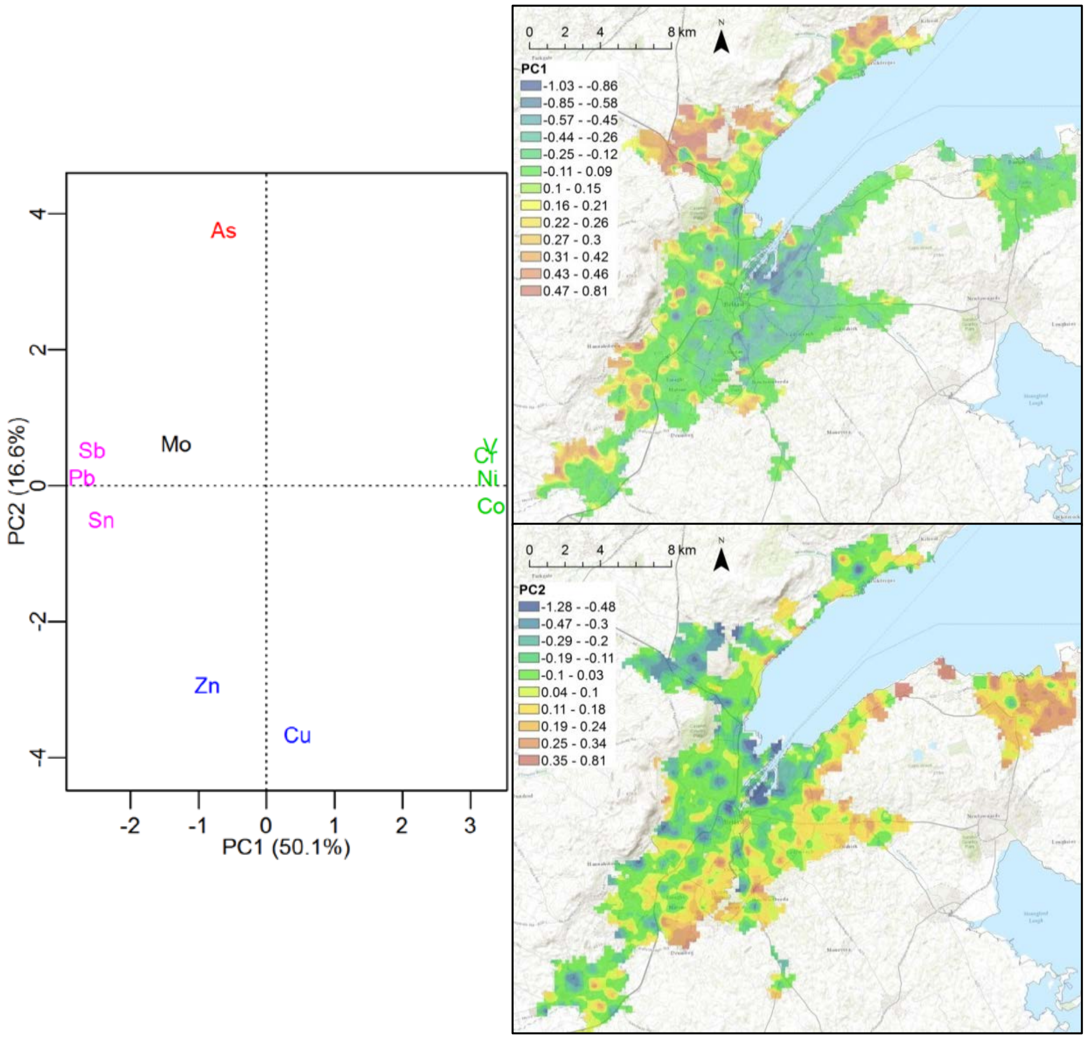


*Figure 4 Results of PCA completed for Belfast using shallow XRF data, with PC1 and PC2 maps*

The first two identified PCs explain 67 % of the total variance within the results of the Belfast PCA. PC1 separates the geogenic controls over Co, V, Cr and Ni from the anthropogenic PTEs. Cobalt, V, Cr and Ni (green on biplot) cluster closely reemphasising their almost identical spatial distributions related to their similar sources. The strong control exerted by the Tertiary basalts over these PTEs is represented in the red areas in the PC1 map (Figure 4).

The PC1 results distinguish Pb, Sb and Sn (pink on biplot) as the most clearly defined anthropogenic PTEs. The map of PC1 shows the anthropogenic group to create a halo effect around the oldest part of the city, with a stronger presence towards the east of the city. The shipbuilding industry has been based within this area of this city for many years, with George Best Belfast City Airport now also located here. This pattern is particularly obvious from the 1901 zone out to the modern Belfast zone, suggesting a long-term pattern of contamination within these soils.

PC2 appears to be explained by different contaminant sources contributing to As (red on biplot), and Cu and Zn (blue on biplot). The widespread As contributions suggest a domestic source such as coal combustion (Duan & Tan 2013), whereas the dark areas related to Cu and Zn suggest point sources across the city centre and also highlight the area of Tertiary basalts in the west of the city. A geogenic contribution to As in areas overlying Silurian greywacke is also possible; although the concentrations of As in this type of bedrock would not be expected to be particularly elevated, we would expect them to be higher than those found over the Tertiary basalts (Young & Donald 2013).

An airborne diffuse source of PTEs aligns well with previous research which demonstrated that characteristic polycyclic aromatic hydrocarbon (PAH) ratios in soils were from an airborne diffuse source from a mixture of biomass, solid fuel and fossil fuel combustion (Doherty et al. 2015).

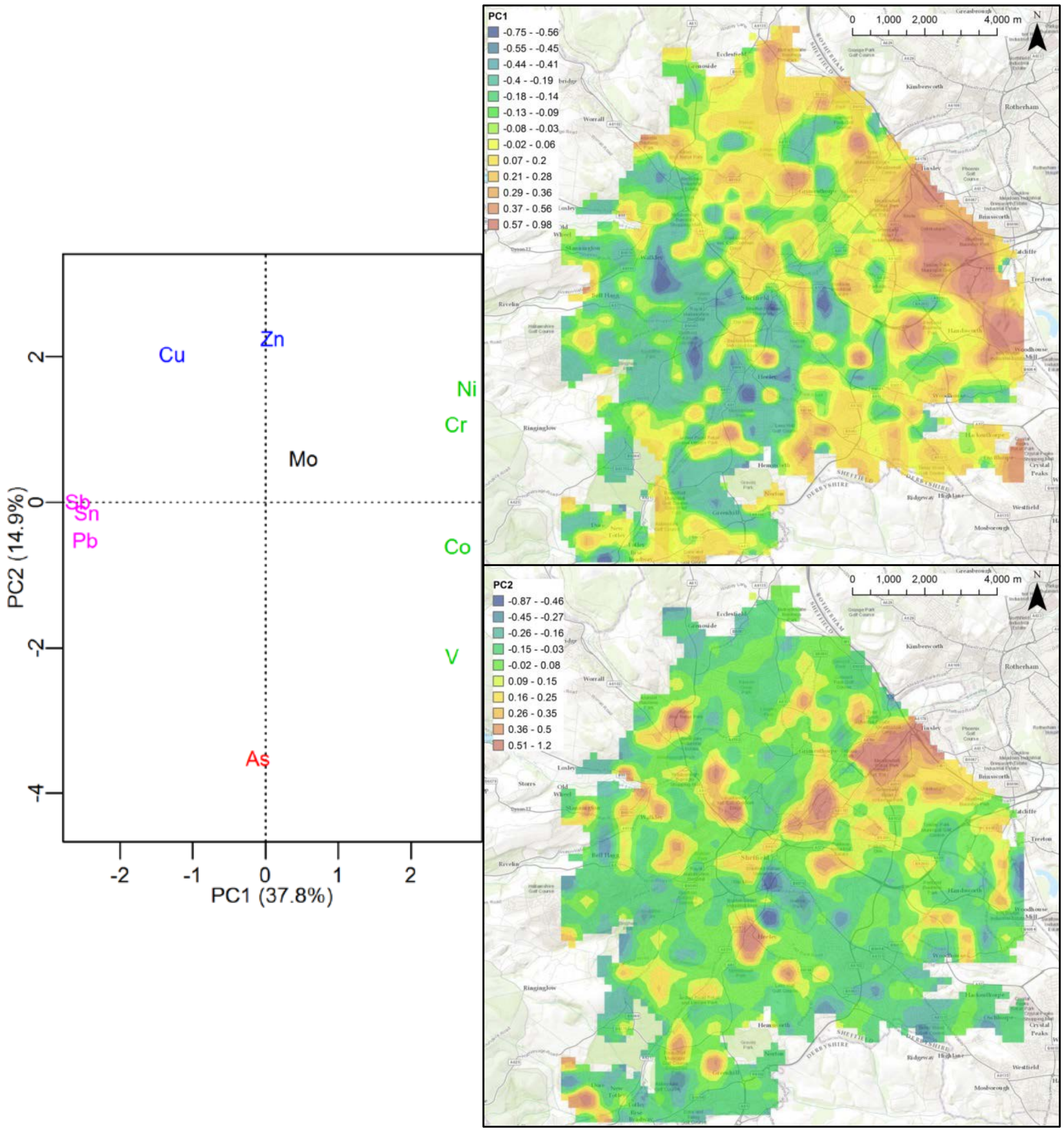


*Figure 5 Results of PCA completed for Sheffield using shallow XRF data, with PC1 and PC2 maps*

The PCA outputs for Sheffield are similar to those for Belfast (Figure 5). In total PC1 and PC2 account for 53% of the total variance within the dataset. This is 14% less than the PCA completed for Belfast highlighting that there may be additional environmental and geochemical parameters affecting variance, for example current land use and activity, in Sheffield.

Within PC1 the results again clearly separate Ni, Cr, Co and V (green) from Sb, Pb and Sn (pink). In contrast to the Belfast results, Ni, Cr, Co and V are distributed much more widely on the biplot (Figure 5), demonstrating that they are still anticipated to be from the same or similar sources, though not as closely correlated in Sheffield as in Belfast. The map for PC1 demonstrates that Ni, Cr, Co and V exert the greatest control over a large area along the north-east boundary of the study area where many of Sheffield's iron and steel works were/are found (as shown in red and orange in the PC1 map (Figure 5)). The elevated concentrations could be related to the industrial use (in these various factories) of the coal which also occurs naturally in this area.

The darker blue colours on the PC1 map (Figure 5) represent the areas controlled by Pb, Sb and Sn, and similarly to Belfast these PTEs are shown to form a halo around the oldest area of the city. The widespread nature of these PTEs on their total concentration maps (Supplementary Information 1) suggests an atmospheric deposition anthropogenic source.

For Sheffield, PC2 separates As (red), from Cu and Zn (blue). In contrast to the Belfast PCA results, Ni and Cr also seem to have an influence over PC2, falling towards the same side as Zn and Cu on the biplot. The blue areas on the PC2 map are related to elevated concentrations of As, suggesting a particular point source of As in these areas. The red areas are related to Zn, Cu, Ni and Cr and fall in a similar, though smaller, area to the north-east of the location identified in PC1. As shown in Figure 6, PC2 seems to be influenced by the geogenic presence of coal with many of the red areas on the map aligning with coal outcrops in Sheffield. Therefore, PC2 seems to identify a geogenic or mining contribution to PTE concentrations (Zn, Cu, Ni and Cr) in Sheffield in the form of coal outcrops.

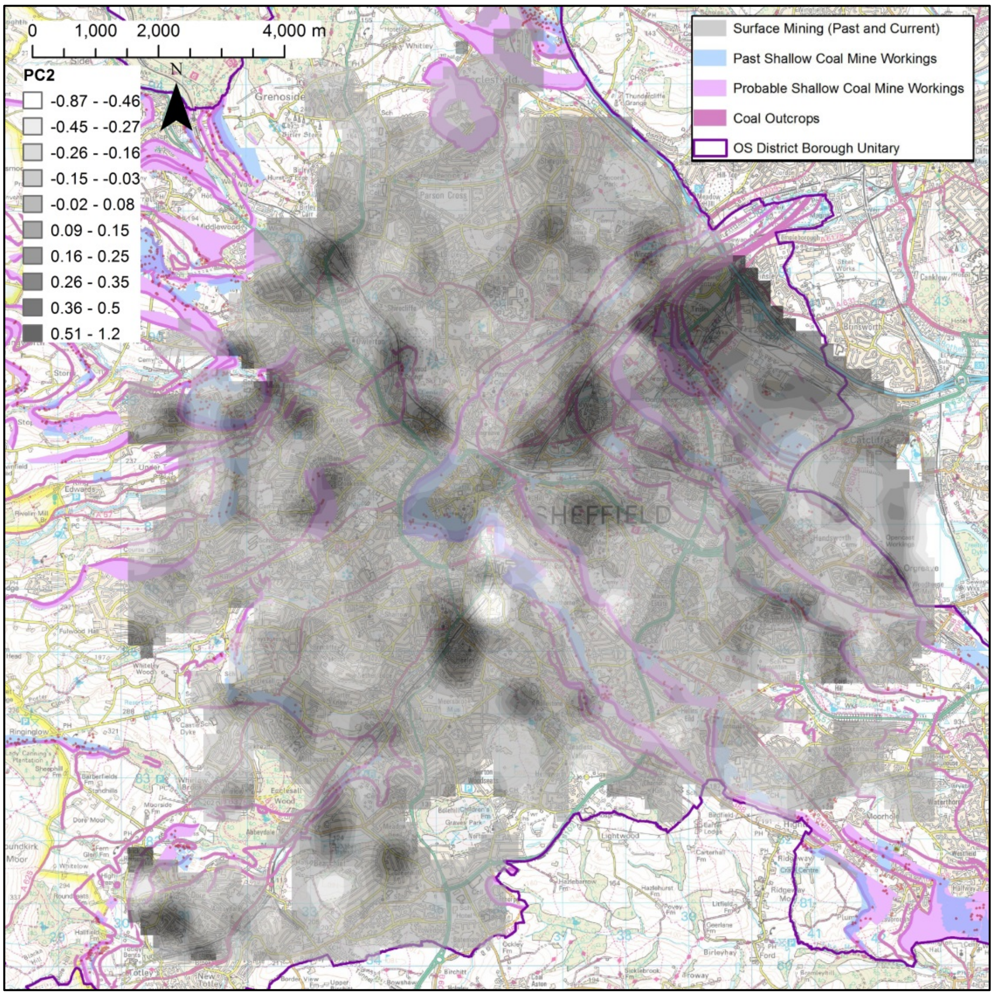


*Figure 6 Map of PC2 from PCA completed for Sheffield using shallow XRF data compared with the specific coal mining legacy plan for Sheffield created by the Coal Authority* (The Coal Authority 2015)

## 3.4 Pollution indices

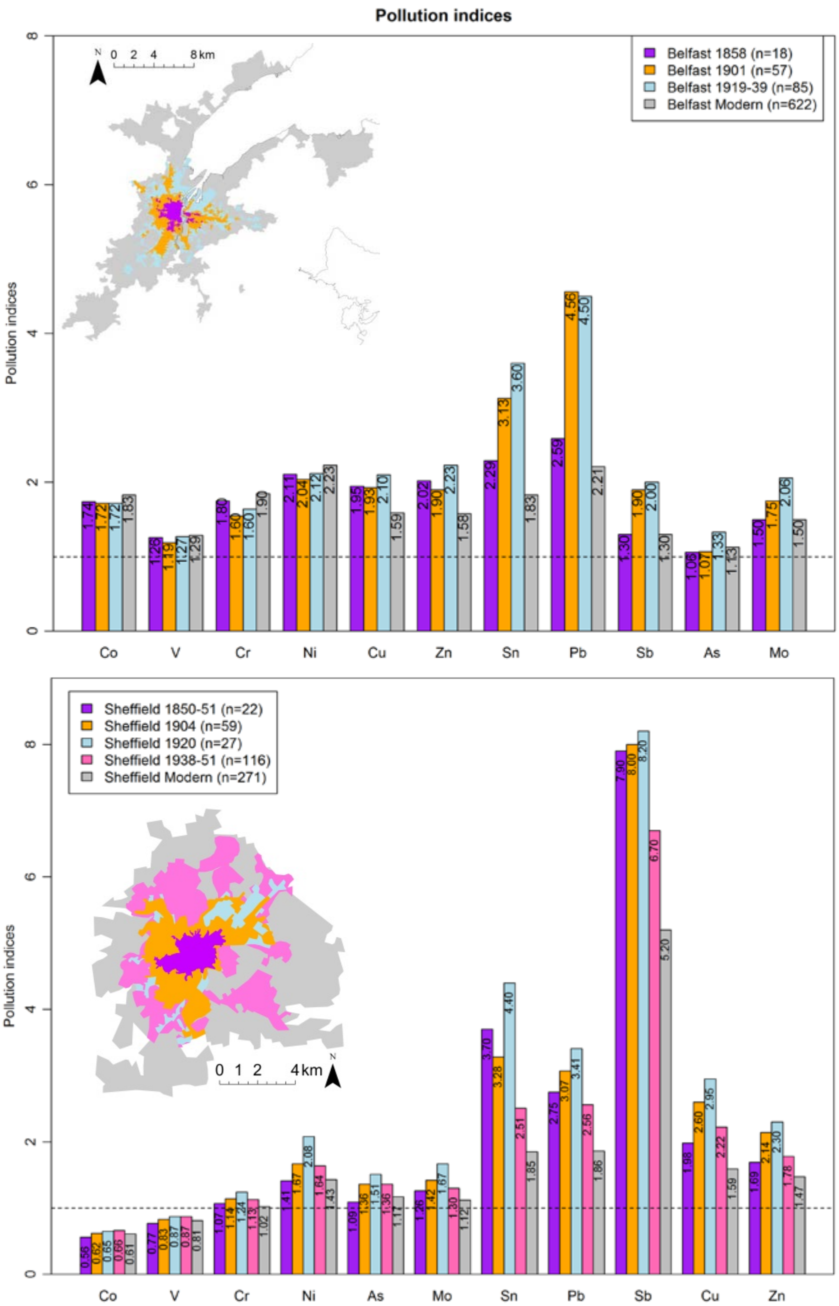


*Figure 7 Bar chart showing the PI for the PTEs within each of Belfast's and Sheffield's development zones (dashed line where PI = 1)*

Pollution indices are used to assess the enrichment of the PTEs in the different historical development zones of both cities. In Belfast, the PIs are calculated using rural median concentrations, thereby demonstrating enrichment of these PTEs in the different Belfast zones compared to rural Northern Ireland. The PIs suggest that certain groups of PTEs are related to different development zones of Belfast (Figure 7).

For Co, V, Cr and Ni the highest PI is within the modern Belfast zone, due to the fact that a greater proportion of the modern Belfast area overlies basalts, which have a significant control over their concentrations, than in the remainder of the development zones. These PIs are much reduced when the median concentration of the PTEs in soils overlying areas of Tertiary basalts replaces the rural median concentration (Supplementary Information 2). For example, the PIs for V, Co, Cr and Ni within the modern development zone reduce from 1.29, 1.83, 1.85 and 2.23 to 0.59, 0.69, 0.62 and 0.66 respectively.

Pb concentrations are at their highest within the 1901 Belfast zone, although the PI is only slightly higher at 4.56 than that within the 1919-1939 zone (4.50). For Cu, Zn, Sn, Sb, As and Mo the highest PIs are located within the 1919-1939 zone suggesting this development zone is the most enriched for the anthropogenically controlled PTEs. However, these PTEs do differ across historical zones: the second highest PI (Cu and Zn – constituents of brass) falls within the 1858 zone, Sn, Sb and Mo have an elevated PI within the 1901 zone, and As within the modern Belfast zone. This suggests that rapid growth associated with development of heavy industry in Belfast between 1901 and 1919-1939 may be responsible for the elevated concentrations of a number of PTEs. However, contamination from all of these PTEs is likely to have begun before this period of time.

Overall, Pb and Sn have the highest pollution indices suggesting the biggest anthropogenic enrichment of these PTEs in the study area. These are also two of the oldest and best recognised urban contaminants. Although Sb and Mo also show anthropogenic enrichment, it is at a much lower level with PIs ranging between 1.30 and 2.00 for Sb and 1.50 and 2.06 for Mo. Geogenic contributions to As in the form of natural mineralisation (McIlwaine et al. 2014) in rural areas of Northern Ireland probably results in lower PIs for As.

As opposed to the PIs calculated for Belfast, the PIs for Sheffield are calculated using a rural median solely from the Carboniferous Coal Measures, rather than the median concentrations from the entire surrounding rural area. The results therefore tell us about enrichment of these PTEs in the urban area compared to one type of bedrock geology, rather than enrichment compared to the rural area as a whole.

The PIs calculated for Sheffield are consistently highest within the 1920 zone. This suggests that industry in Sheffield between 1904 and 1920 may be responsible for the most elevated PTE concentrations in soils. For Sn, Pb and Sb the PIs fall considerably within the 1938-1951 zone, below all three of the previous historical zones. The PIs are lowest for Cr, Mo, Sn, Pb, Sb, Cu and Zn in the modern zone, and lowest for Co, V, Ni and As in the 1850-1851 zone.

Lead, Sb and Sn show the greatest accumulation in Sheffield's soils (highest PIs), followed by Cu and Zn. The PIs for Sb are high due to its low median concentration in the Carboniferous Coal Measures (0.5 mg/kg). Cobalt and V are depleted in Sheffield's soils when compared to the rural median, while Cr, Ni, As and Mo show minimal accumulation.

## 3.5 PTE typical threshold values

Within the Belfast study area, four PTEs (Co, Cr, Ni and V) have been clearly identified as elements that are geogenically controlled. The remainder of the PTEs in Belfast (As, Cu, Mo, Pb, Sb, Sn and Zn) are dominated by anthropogenic inputs. In Sheffield, all the PTEs show an anthropogenic influence. From these results, different methods for calculating background concentrations have been applied to assess the concentrations of the anthropogenically controlled PTEs within the city development zones, which have been shown to act as urban subdomains. These background values indicate what a 'typical' or background concentration of these PTEs would be within the defined domain; aiming to differentiate between concentrations related to geogenic and diffuse anthropogenic sources, and concentrations generated by point sources.

*Table 1 Summary of background values calculated for the PTEs in Belfast regarded as having some anthropogenic input within the separate development zones via the ULBL, Median + 2MAD and NBC methods (BC = box-cox transformation, L = log transformation and E = empirical)*

| | 1858 | | | 1901 | | | 1919-1939 | | | Modern | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC |
| As | 18 | 13 | - | 21 | 15 | BC 37 | 26 | 17 | E 52 | 19 | 14 | L 21 |
| Cu | 120 | 95 | - | 160 | 100 | BC 210 | 200 | 120 | BC 640 | 120 | 80 | L 130 |
| Mo | 2.8 | 1.8 | - | 3.5 | 2.4 | L 4.7 | 5.1 | 3.0 | E 18 | 2.7 | 1.8 | L 3.1 |
| Pb | 190 | 140 | - | 430 | 280 | L 490 | 620 | 270 | BC 1300 | 200 | 120 | L 260 |
| Sb | 2.7 | 1.9 | - | 4.3 | 3.1 | L 10 | 7.2 | 3.7 | BC 33 | 3.0 | 2.1 | L 4.3 |
| Sn | 16 | 11 | - | 20 | 14 | L 33 | 51 | 18 | BC 1000 | 14 | 7.7 | BC 24 |
| Zn | 240 | 210 | - | 310 | 220 | L 470 | 510 | 290 | BC 2100 | 240 | 170 | L 290 |

Normal background concentrations cannot be calculated for the 1858 Belfast domain as there are only 18 samples available within this area. With the exception of the Median + 2MAD method for Pb, all the background values calculated are highest within the 1919-1939 zone. This would be expected for all the PTEs apart from Pb, where the calculated PI is highest within the 1901 zone (Figure 7) suggesting the greatest enrichment of Pb in this zone. Upon further investigation this was found to be related to the distribution of the data. As the PI is based on the median of the dataset it is less affected by skew than the background value calculations via the ULBL and NBC methods; the 1919-1939 Pb data is more highly skewed than the 1901 Pb data. This suggests a more homogenous source of Pb, such as atmospheric deposition, within the 1901 zone, whereas the 1919-1939 zone is possibly witnessing atmospheric deposition as well as more independent point sources of Pb. This is possibly associated with increased development in the east of the city within this 1919-1939 zone, as identified and discussed within the PC1 results (Section 3.3).

*Table 2 Summary of background values calculated for the PTEs in Sheffield regarded as having some anthropogenic input within the separate development zones via the ULBL, Median + 2MAD and NBC methods (BC = box-cox transformation, L = log transformation, N = no transformation and E = empirical)*

| | 1850-1851 | | | 1904 | | | 1920 | | | 1938-1951 | | | Modern | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC | ULBL | M +2MAD | NBC |
| As | 62 | 34 | - | 70 | 44 | L 77 | 54 | 39 | - | 50 | 38 | L 61 | 41 | 31 | L 50 |
| Co | 31 | 22 | - | 29 | 23 | L 40 | 36 | 24 | - | 29 | 23 | E 41 | 29 | 23 | E 28 |
| Cr | 220 | 130 | - | 190 | 130 | E 780 | 270 | 140 | - | 190 | 130 | E 600 | 160 | 120 | E 370 |
| Cu | 210 | 110 | - | 300 | 160 | L 590 | 250 | 150 | - | 180 | 120 | L 260 | 120 | 80 | BC 200 |
| Mo | 12 | 7.4 | - | 12 | 8.3 | L 23 | 15 | 8.3 | - | 9.1 | 6.7 | BC 19 | 8.3 | 5.9 | L 13 |
| Ni | 97 | 57 | - | 82 | 62 | BC 240 | 120 | 82 | - | 75 | 58 | BC 140 | 68 | 51 | L 89 |
| Pb | 940 | 440 | - | 930 | 530 | L 1400 | 800 | 460 | - | 560 | 360 | L 790 | 370 | 250 | L 600 |
| Sb | 16 | 8.0 | - | 14 | 8.2 | L 20 | 12 | 6.7 | - | 11 | 6.2 | BC 22 | 7.2 | 4.6 | L 9.2 |
| Sn | 100 | 47 | - | 77 | 45 | L 170 | 69 | 48 | - | 46 | 28 | L 64 | 33 | 21 | L 46 |
| V | 130 | 110 | - | 140 | 110 | N 140 | 130 | 110 | - | 140 | 110 | E 150 | 140 | 110 | L 130 |
| Zn | 420 | 290 | - | 620 | 400 | L 920 | 520 | 380 | - | 420 | 280 | BC 880 | 330 | 230 | L 390 |

NBCs cannot be calculated for the 1850-1851 or 1920 Sheffield domain as there are only 22 and 27 samples available within these areas respectively. The domains with the most elevated background values are reasonably consistent across the ULBL and Median + 2MAD calculations. Where these differences occur it is likely to be related to the distribution of the data, as explained for Pb in the Belfast results.

The main differences between the ULBL and Median + 2MAD concentrations are noted for Pb, Sb and Sn. The highest concentrations via the ULBL method are in the 1850-1851 zone for all three PTEs, whereas via the Median + 2MAD method the highest concentrations are in the 1904 zone for Pb and Sb and the 1920 zone for Sn. In conjunction with the spatial distribution previously discussed, the most elevated concentrations of all of these PTEs seem to form a halo surrounding the oldest part of the city, suggesting a stronger alignment with the ULBLs than the Median + 2MAD values.

The ULBL method highlights the highest concentrations for Co, Cr, Mo and Ni in the 1920 zone, while the Median + 2MAD method has the highest concentrations for Co, Cr, Mo, Ni, Sn and V in this zone. The concentrations of V do not vary much across the development zones for any the methods used to calculate background values, suggesting this PTE is not influenced by different periods of historical development to the same extent as the other PTEs.

As can be seen in Supplementary Information 3, the most conservative background concentrations (i.e. the lowest concentrations) are calculated by the Median + 2MAD method, while the least conservative (i.e. the highest concentrations) are calculated via the NBC method, for all the PTEs except V. This is true for the majority of the background concentrations calculated, not just for Sheffield's modern zone.

Rothwell & Cooke (2015) suggested the median + 2MAD method for use because it consistently calculated the most conservative background concentrations in their study in Gateshead. Although this is a sensible precaution from a risk perspective, it may not be realistic to state that further investigation may be required at 19% of sampled sites within Sheffield's modern zone where the concentration of Pb is above the calculated Median + 2MAD value. The ULBLs provide concentrations between the most conservative Median + 2MAD concentrations and the least conservative NBCs; they may therefore be more appropriate values to use in gaining an understanding of background concentrations of different PTEs in these studies. In addition, the NBC methodology is only applicable to domains with more than 30 samples and so it can't be applied in a number of the urban subdomains identified.

The ULBL methodology has previously been applied on a Northern Ireland regional scale to generate TTVs (McIlwaine et al. 2014). Although only compared with NBCs at this regional level, it was still identified as the most appropriate method for calculating background values.

### 3.5.1 Comparison with generic assessment criteria

Comparisons can be drawn between the recommended background values calculated via the ULBL method and UK generic assessment criteria such as suitable 4 use levels (S4ULs) and provisional category 4 screening levels (pC4SLs). These different criteria vary slightly in their definition, and are available for different PTEs. Suitable 4 use levels are based on the same level of risk as soil guideline values (SGVs) i.e. minimal or tolerable risk. However, S4ULs were generated using an updated exposure model and are available for additional land uses (public open spaces) (Nathanail et al. 2015). Provisional C4SLs have been created to support Defra's statutory guidance for Part 2A of the Environmental Protection Act. The guidance stated that where there is no risk that land poses a significant possibility of significant harm, or the level of risk is low, the category 4 classification should be used. At the other extreme, category 1 encompasses areas where the risk of the land posing a significant possibility of significant harm is unacceptably high (CL:AIRE 2014a). Therefore the key difference between S4ULs and C4SLs is the level of risk they consider; S4ULs are guidelines considering a level that is tolerable or posing a minimal risk to human health whereas C4SLs describe a higher level of risk which can still be considered low enough to allow category 4 land classification (CL:AIRE 2014a). The available S4ULs and pC4SLs are provided in Supplementary Information 4.

All the calculated background values for As in Belfast fall below the lowest SGV (32 mg/kg), S4UL (37 mg/kg) and pC4SL (37 mg/kg), while all the ULBLs for Sheffield fall above these values. The highest ULBL for Sheffield (70 mg/kg), calculated for the 1904 zone, is higher than the residential and allotment SGVs, pC4SLs and S4ULs (Supplementary Information 4). For Cu and Zn all the ULBLs fall below the lowest available S4ULs (520 mg/kg and 620 mg/kg for Cu and Zn respectively) for both Belfast and Sheffield, with the exception of the 1904 ULBL for Zn which is equal to the allotment S4UL of 620 mg/kg. The comparison for Cr depends upon its speciation; the ULBLs calculated within Sheffield are all lower than the most conservative S4UL for Cr(III), but vast exceedances are obvious across the development zones for Cr(VI). The ULBLs calculated for Ni in Sheffield are all within the most conservative S4UL for Ni (130 mg/kg for residential areas). Exceedances of the most conservative S4UL for V (91 mg/kg for allotments) are shown for all the ULBLs calculated for Sheffield.

The ULBLs calculated for Pb are high, especially in the 1919-1939 development zone for Belfast (620 mg/kg) and the 1850-1851 zone for Sheffield (940 mg/kg). This is of concern as Pb is known to be a non-threshold toxin i.e. no minimal risk level has been identified (Palmer et al. 2015). The lowest published pC4SL for Pb is 34 mg/kg, identified for allotment land use (CL:AIRE 2014b). The ULBLs calculated are 5.6, 12.6, 18.2 and 5.9 times greater than this pC4SL within the 1858, 1901, 1919-1939 and modern development zones respectively in Belfast, and 27.6, 27.4, 23.5, 16.5 and 10.9 times greater in the 1850-1851, 1904, 1920, 1938-1951 and modern development zones in Sheffield. These calculated ULBLs are elevated when compared to the pC4SLs for Pb, and for sample locations where the ULBLs are exceeded in particular, further investigation may be required to examine the potential risk posed.

# 4 Conclusions

The scale of PTE concentration data available within the study area considered allowed for a thorough examination of the effects of historical development on soil PTE concentrations. Clear groups of PTEs were identified within the study area investigated via depth ratios, a range of multivariate statistical techniques and PIs. In particular, depth ratios proved to be a useful technique for identifying controlling sources in this urban environment. The concentrations in shallow soils were found to be controlled to a greater extent by anthropogenic influences than concentrations in deeper soils, which remain controlled by principally geogenic processes. Controlling sources and links between historical development and PTE concentrations were identified, suggesting the investigative methodology employed within this research may be useful for application within other urban environments.

The depth ratio boxplots suggested the highest levels of anthropogenic input for Pb, Sb and Sn followed by As and Mo in Belfast. Geogenic inputs were found to control the concentrations of Ni, Co, Cr and V, while Cu and Zn were influenced by both anthropogenic and geogenic inputs. The marked similarity in the spatial distributions of the geogenically controlled PTEs (Co, V, Cr and Ni) clearly demonstrates the control that the Tertiary basalts have over their concentrations, and suggests similar point sources of anthropogenic contributions perhaps due to importing of 'clean' topsoil, which originally overlaid the Antrim basalts, in these areas. PTEs under predominantly anthropogenic sources in Belfast can be split into three groups; 1) Sn, Pb and Sb, 2) Cu and Zn and 3) As and Mo. Cu and Zn receive some geogenic contribution to their concentrations from the Tertiary basalts but similar anthropogenic contributions are also obvious. Increasing anthropogenic contributions to both As and Mo see them grouped similarly, with Pb, Sb and Sn noted for the greatest anthropogenic contribution.

All of the PTEs investigated were found to be under some anthropogenic influence in Sheffield. Nickel, Co, Cr and V were found to align well in a large area along the north-east boundary of Sheffield where they were shown to occur at elevated concentrations. This pattern is thought to be related to the various industrial land uses located in this part of the city; many of Sheffield's iron and steel works were/are found here. The elevated concentrations could be related to the industrial use (in these various factories) of the coal which also occurs naturally in this area. Similarly to Belfast, Pb, Sb and Sn were shown to form a halo around the oldest area of the city. The widespread nature of these PTEs on their total concentration maps suggests anthropogenic atmospheric source deposition. A geogenic contribution to Zn, Cu, Ni and Cr concentrations was identified in the form of coal outcrops in Sheffield.

The relationship between historical development and differing PTEs is a novel finding from this research. This suggests that PTEs have the potential for use as 'urbanisation tracers' as different PTEs have been shown to be associated with different historical anthropogenic sources.

Background values were calculated for the PTEs deemed to have some form of anthropogenic input (via the PCA) within each of the city's development zones. These PTEs demonstrated varying historical sources within the development zones which result in varying background values. The background values for Belfast were generally highest in the 1919-1939 development zone, suggesting that soil contamination was at its greatest in Belfast between 1901 and 1939. This coincides with when shipbuilding and its associated industries, such as iron and steel foundries, were at their zenith. The background values calculated for Sheffield varied more widely across the different development zones considered. The ULBL method was determined to be the most appropriate for calculating background values within this research, and the ULBLs were compared to SGVs, S4ULs and pC4SLs. A number of exceedances of these generic assessment criteria were noted for various PTEs in various development zones. This demonstrates that the potential risk associated with PTEs in urban environments may require further assessment. The development of urban subdomains refines the areas that need to be considered for additional investigation.

Current statutory guidance for Part 2A of the Environmental Protection Act states that widespread geogenic or diffuse anthropogenic pollution in soil should not be regarded as contaminated land unless other evidence demonstrates that it poses a risk. These findings establish that PTE concentrations associated with geogenic and diffuse anthropogenic contamination are elevated to levels that do have the potential to pose risk to human health. However, as with anthropogenic point sources, a regional assessment of PTE bioaccessibility would need to be completed to assess the level of risk posed. These findings potentially suggest that a relevant legislative regime that ensures geogenic and diffuse anthropogenic contamination are effectively dealt with may be required.

## Acknowledgements

This research has been funded by the Annual Scholarship of the Society of Brownfield Risk Assessment (SoBRA). The authors thank SoBRA for this funding and for recognising the significance of this work. Alex Donald of the Geological Survey of Northern Ireland (GSNI) is thanked for arranging access to the Tellus data. The Tellus project was funded by the Department of Enterprise Trade and Investment and by the Rural Development Programme through the Northern Ireland Programme for building sustainable prosperity. Dr Doherty and Dr Cox were supported by the REMEDIATE (Improved decision-making in contaminated land site investigation and risk assessment) Marie-Curie Innovation Training Network. The network has received funding from the European Union's Horizon 2020 Programme for research, technological development and demonstration under grant agreement n. 643087. The authors declare that they have no conflict of interest.

# References


1:25000 maps of Great Britain, 1953. *SK38, SK39, SK48 and SK49*,

Aitchison, J., 1982. The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society*, 44(2), pp.139–177.

Ajmone-Marsan, F. & Biasioli, M., 2010. Trace Elements in Soils of Urban Areas. *Water, Air, & Soil Pollution*, 213, pp.121–143. Available at: http://link.springer.com/10.1007/s11270-010-0372-6 [Accessed December 3, 2014].

Andersson, M., Ottesen, R.T. & Langedal, M., 2010. Geochemistry of urban surface soils - Monitoring in Trondheim, Norway. *Geoderma*, 156, pp.112–118. Available at: http://dx.doi.org/10.1016/j.geoderma.2010.02.005.

Argyraki, A. & Kelepertzis, E., 2014. Urban soil geochemistry in Athens, Greece: The importance of local geology in controlling the distribution of potentially harmful trace elements. *The Science of the total environment*, 482-483, pp.366–377. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24662205 [Accessed September 19, 2014].

Astel, A., Tsakovski, S., Barbieri, P. & Simeonov, V., 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. *Water research*, 41, pp.4566–4578. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17632213 [Accessed February 11, 2013].

Barsby, A., McKinley, J.M., Ofterdinger, U., Young, M., Cave, M.R. & Wragg, J., 2012. Bioaccessibility of trace elements in soils in Northern Ireland. *The Science of the total environment*, 433, pp.398–417. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22819891 [Accessed November 29, 2012].

Bartholomew's "Half Inch Maps" of England and Wales, 1904. *Sheet 9 - Sheffield*,

Bartholomew's Revised "Half Inch Maps," 1920. *Sheet 9 - Sheffield*,

Beckett, J.C. & Glasscock, R.E. eds., 1967. *Belfast: Original & growth of an industrial city*, British Broadcasting Corporation.

Biasioli, M., Barberis, R. & Ajmone-Marsan, F., 2006. The influence of a large city on some soil properties and metals content. *The Science of the total environment*, 356, pp.154–164. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15941578 [Accessed December 18, 2014].

Candeias, C., Ferreira da Silva, E., Salgueiro, A.R., Pereira, H.G., Reis, A.P., Patinha, C., Matos, J.X. & Ávila, P.H., 2011. The use of multivariate statistical analysis of geochemical data for assessing the spatial distribution of soil contamination by potentially toxic elements in the Aljustrel mining area (Iberian Pyrite Belt, Portugal). *Environmental Earth Sciences*, 62, pp.1461–1479. Available at: http://www.springerlink.com/index/10.1007/s12665-010-0631-2 [Accessed November 22, 2012].

Carrero, J.A., Arrizabalaga, I., Bustamante, J., Goienaga, N., Arana, G. & Madariaga, J.M., 2013. Diagnosing the traffic impact on roadside soils through a multianalytical data analysis of the concentration profiles of traffic-related elements. *The Science of the total environment*, 458-460, pp.427–434. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23685368 [Accessed February 4, 2014].

Cave, M.R., Johnson, C.C., Ander, E. & Palumbo-Roe, B., 2012. Methodology for the determination of normal background contaminant concentrations in English soils. *British Geological Survey Commissioned Report*, CR/12/003, p.42pp.

Chiprés, J.A., de la Calleja, A., Tellez, J.I., Jiménez, F., Cruz, C., Guerrero, E.G., Castro, J., Monroy, M.G. & Salinas, J.C., 2009. Geochemistry of soils along a transect from Central Mexico to the Pacific Coast: A pilot study for continental-scale geochemical mapping. *Applied Geochemistry*, 24, pp.1416–1428. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0883292709001231 [Accessed February 20, 2013].

Cole, S. & Jeffries, J., 2009. *Using Soil Guideline Values*, Environment Agency, Bristol. Available at: http://www.environment-agency.gov.uk/static/documents/Research/SCHO0309BPQM-e-e.pdf.

Contaminated Land: Applications in Real Environments (CL:AIRE), 2014a. SP1010 – Development of Category 4 Screening Levels for Assessment of Land Affected by Contamination. *Final Project Report (Revision 2)*, (September), pp.1–111.

Contaminated Land: Applications in Real Environments (CL:AIRE), 2014b. *SP1010 – Development of Category 4 Screening Levels for Assessment of Land Affected by Contamination: Appendix H Provisional C4SLs for Lead*, London, Contaminated Land: Applications in Real Environments.

Cox, S.F., Chelliah, M.C.M., McKinley, J.M., Palmer, S., Ofterdinger, U., Young, M.E., Cave, M.R. & Wragg, J., 2013. The importance of solid-phase distribution on the oral bioaccessibility of Ni and Cr in soils overlying Palaeogene basalt lavas, Northern Ireland. *Environmental Geochemistry and Health*, 35, pp.553–567. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23821222 [Accessed August 24, 2013].

Crawford, W.H., 1986. *Industries of the North one hundred years ago: Industrial and Commercial Life in the North of Ireland 1888-91* 2nd ed., Belfast: Universities Press Ltd.

Dao, L., Morrison, L., Zhang, H. & Zhang, C., 2014. Influences of traffic on Pb, Cu and Zn concentrations in roadside soils of an urban park in Dublin, Ireland. *Environmental geochemistry and health*, 36, pp.333–343. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23828236 [Accessed April 29, 2014].

Dirks, K.N., Hay, J.E., Stow, C.D. & Harris, D., 1998. High-resolution studies of rainfall on Norfolk Island Part II: Interpolation of rainfall data. *Journal of Hydrology*, 208, pp.187–193.

Doherty, R., McIlwaine, R., McAnallen, L. & Cox, S., 2015. Source Apportionment, Fugacity Modelling and Comparison of Background Concentration Methods of Soil PAHs in an Urban Setting. In O'Sullivan & Megson, eds. *International Network of Environmental Forensics Conference*. Cambridge: RSC.

Duan, J. & Tan, J., 2013. Atmospheric heavy metals and Arsenic in China: Situation, sources and control policies. *Atmospheric Environment*, 74, pp.93–101. Available at: http://dx.doi.org/10.1016/j.atmosenv.2013.03.031.

European Environment Agency, 2012. Corine Land Cover 2006 seamless vector data. , p.Version 16 (04/2012). Available at: http://www.eea.europa.eu/data-and-maps/data/clc-2006-vector-data-version-2#tab-additional-information.

Freestone, S.E., O'Donnell, K.E. & Brown, S.E., 2004. Geochemical baseline data for the urban area of Sheffield. *British Geological Survey Internal Report*, IR/02/084, p.71. Available at: http://nora.nerc.ac.uk/7019/.

Frentiu, T., Pintican, B.P., Butaciu, S., Mihaltan, A.I., Ponta, M. & Frentiu, M., 2013. Determination, speciation and distribution of mercury in soil in the surroundings of a former chlor-alkali plant: assessment of sequential extraction procedure and analytical technique. *Chemistry Central Journal*, 7, pp.1–14. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24252185.

Galán, E., Fernández-Caliani, J., González, I., Aparicio, P. & Romero, A., 2008. Influence of geological setting on geochemical baselines of trace elements in soils. Application to soils of South–West Spain. *Journal of Geochemical Exploration*, 98, pp.89–106. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0375674208000071 [Accessed March 26, 2013].

Gilbertson, D., Grattan, J., Cressey, M. & Pyatt, F., 1997. An air-pollution history of metallurgical innovation in iron and steel-making: a geochemical archive of Sheffield. *Water, Air and Soil Pollution1*, 100, pp.327–341.

Glennon, M.M., Harris, P., Ottesen, R.T., Scanlon, R.P. & O'Connor, P.J., 2014. The Dublin SURGE Project: geochemical baseline for heavy metals in topsoils and spatial correlation with historical industry in Dublin, Ireland. *Environmental geochemistry and health*, 36(2), pp.235–54. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3938859&tool=pmcentrez&rendertype=abstract [Accessed April 29, 2014].

Golden, N., Morrison, L., Gibson, P.J., Potito, A.P. & Zhang, C., 2015. Spatial patterns of metal contamination and magnetic susceptibility of soils at an urban bonfire site. *Applied Geochemistry*, 52, pp.86–96. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0883292714002662.

Green, K.A., Caven, S. & Lister, T.R., 2010. Tellus Soil Geochemistry - Quality Assessment and Map Production of ICP data. *British Geological Survey Internal Report*, IR/11/01, p.142.

Herting, G., Goidanich, S., Odnevall Wallinder, I. & Leygraf, C., 2008. Corrosion-induced release of Cu and Zn into rainwater from brass, bronze and their pure metals. A 2-year field study. *Environmental Monitoring and Assessment*, 144(1-3), pp.455–461.

Jarva, J., Tarvainen, T., Reinikainen, J. & Eklund, M., 2010. TAPIR - Finnish national geochemical baseline database. *Science of the Total Environment*, 408, pp.4385–4395.

Johnson, C.C., 2005. 2005 G-BASE Field Procedures Manual. *British Geological Survey Internal Report*, IR/05/097, p.130.

Johnson, C.C., Ander, E., Cave, M.R. & Palumbo-Roe, B., 2012. Normal background concentrations (NBCs) of contaminants in English soils: Final project report. *British Geological Survey Commissioned Report*, CR/12/035, p.40pp.

Johnson, C.C. & Ander, E.L., 2008. Urban geochemical mapping studies: how and why we do them. *Environmental geochemistry and health*, 30, pp.511–530. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18563589 [Accessed February 11, 2014].

Johnson, C.C., Demetirades, A., Locutura, J. & Ottesen, R.T. eds., 2011. *Mapping the Chemical Environment of Urban Areas*, John Wiley & Sons, Ltd.

Karlsson, T., Persson, P. & Skyllberg, U., 2006. Complexation of Copper(II) in Organic Soils and in Dissolved Organic Matter − EXAFS Evidence for Chelate Ring Structures. *Environmental Science & Technology*, 40(8), pp.2623–2638.

Kelepertzis, E. & Argyraki, A., 2015. Geochemical associations for evaluating the availability of potentially harmful elements in urban soils: Lessons learnt from Athens, Greece. *Applied Geochemistry*. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0883292715000839.

Knights, K. V, 2007. A report on the Tellus urban field campaigns of Belfast Metropolitan areas and Londonderry. *British Geological Survey Commissioned Report*, CR/07/006N, p.19.

Land and Property Services, 1858. *Ordnance Survey 6 Inch County Maps*, © Crown copyright and database rights MOU203.

Land and Property Services, 1901. *Ordnance Survey 6 Inch County Maps*, © Crown copyright and database rights MOU203.

Land and Property Services, 1919. *Ordnance Survey 6 Inch County Maps*, © Crown copyright and database rights MOU203.

Marchant, B.P., Tye, A.M. & Rawlins, B.G., 2011. The assessment of point-source and diffuse soil metal pollution using robust geostatistical methods: a case study in Swansea (Wales, UK). *European Journal of Soil Science*, 62, pp.346–358. Available at: http://doi.wiley.com/10.1111/j.1365-2389.2011.01373.x [Accessed December 15, 2014].

McIlwaine, R., Cox, S.F. & Doherty, R., 2015. When are total concentrations not total? Factors affecting geochemical analytical techniques for measuring element concentrations in soil. *Environmental Science and Pollution Research*, 22, pp.6364–6371.

McIlwaine, R., Cox, S.F., Doherty, R., Palmer, S., Ofterdinger, U. & McKinley, J., 2014. Comparison of methods used to calculate typical threshold values for potentially toxic elements in soil. *Environmental Geochemistry and Health*, 36(5), pp.953–971.

Mielke, H.W., 1999. Lead in the Inner Cities: Policies to reduce children's exposure to lead may be overlooking a major source of lead in the environment. *American Scientist*, 87, pp.62–73.

Mitchell, W. ed., 2004. *The Geology of Northern Ireland* Second Edi., Belfast: Geological Survey of Northern Ireland.

Nathanail, C.P., McCaffrey, C., Gillett, A.G., Ogden, R.C. & Nathanail, J.F., 2015. *The LQM/CIEH S4ULs for Human Health Risk Assessment*, Nottingham: Land Quality Press. Copyright Land Quality Management Limited reproduced with permission; Publication Number S4UL3374. All rights reserved.

Northern Ireland Statistics and Research Agency, 2014. Northern Ireland Census of Employment 2013. *Labour Market Statistics Bulletin*, (July).

OS Six Inch England and Wales, 1851. *Yorkshire 294 (includes Sheffield)*,

Palmer, S., McIlwaine, R., Ofterdinger, U., Cox, S.F., McKinley, J.M., Doherty, R., Wragg, J. & Cave, M.R., 2015. The effects of lead sources on oral bioaccessibility in soil and implications for contaminated land risk management. *Environmental Pollution*, 198, pp.161–171.

Pawlowsky-Glahn, V. & Egozcue, J.J., 2006. Compositional data and their analysis: an introduction. *Geological Society, London, Special Publications*, 264, pp.1–10. Available at: http://sp.lyellcollection.org/cgi/doi/10.1144/GSL.SP.2006.264.01.01 [Accessed March 6, 2013].

Rawlins, B.G., Lark, R.M., O'Donnell, K.E., Tye, A.M. & Lister, T.R., 2005. The assessment of point and diffuse metal pollution of soils from an urban geochemical survey of

Sheffield, England. *Soil Use and Management*, 21, pp.353–362. Available at: http://doi.wiley.com/10.1079/SUM2005335 [Accessed October 19, 2012].

Reimann, C., Filzmoser, P., Fabian, K., Hron, K., Birke, M., Demetriades, A., Dinelli, E. & Ladenberger, A., 2012. The concept of compositional data analysis in practice--total major element concentrations in agricultural and grazing land soils of Europe. *The Science of the total environment*, 426, pp.196–210. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22503552 [Accessed March 6, 2013].

Reimann, C., Filzmoser, P. & Garrett, R.G., 2005. Background and threshold: critical comparison of methods of determination. *The Science of the total environment*, 346, pp.1–16. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15993678 [Accessed March 12, 2012].

Reimann, C., Filzmoser, P., Garrett, R.G. & Dutter, R., 2008. *Statistical Data Analysis Explained: Applied Environmental Statistics with R*, Chichester: John Wiley & Sons Ltd.

Reimann, C. & Garrett, R.G., 2005. Geochemical background--concept and reality. *The Science of the total environment*, 350, pp.12–27. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15890388 [Accessed October 19, 2012].

Rodrigues, S., Urquhart, G., Hossack, I., Pereira, M.E., Duarte, A.C., Davidson, C., Hursthouse, A., Tucker, P. & Roberston, D., 2009. The influence of anthropogenic and natural geochemical factors on urban soil quality variability: a comparison between Glasgow, UK and Aveiro, Portugal. *Environmental Chemistry Letters*, 7, pp.141–148. Available at: http://link.springer.com/10.1007/s10311-008-0149-y [Accessed January 9, 2015].

Rothwell, K.A. & Cooke, M.P., 2015. A comparison of methods used to calculate normal background concentrations of potentially toxic elements for urban soil. *Science of The Total Environment*, 532, pp.625–634. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0048969715302783.

Royle, S.A., 2007. *Belfast Part II, 1840 to 1900* A. Simms, H. B. Clarke, & R. Gillespie, eds., Royal Irish Academy in association with Belfast City Council: Dublin University Press.

Smyth, D., 2007. Methods used in the Tellus Geochemical Mapping of Northern Ireland. *British Geological Survey Open Report*, OR/07/022, p.90pp.

The Coal Authority, 2015. *Sheffield District (B) - Specific Risk*, Available at: https://www.gov.uk/government/publications/coalfield-plans-sheffield-city-council-area.

Thorton, I., 2009. Chapter 4 Metal Contamination of Soils in Urban Areas. In P. Bullock & P. J. Gregory, eds. *Soils in the Urban Environment*. Blackwell Scientific Publications.

United Nations Department of Economic and Social Affairs Population Division, 2014. *World Urbanization Prospects: The 2014 Revision*, Highlights (ST/ESA/SER.A/352), United Nations.

Ward, J.H., 1963. Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58(301), pp.236–244.

Young, M. & Donald, A. eds., 2013. *A guide to the Tellus data*, Belfast: Geological Survey of Northern Ireland.

Zhang, C., Jordan, C. & Higgins, A., 2007. Using neighbourhood statistics and GIS to quantify and visualize spatial variation in geochemical variables: An example using Ni concentrations in the topsoils of Northern Ireland. *Geoderma*, 137, pp.466–476. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0016706106003053 [Accessed October 19, 2012].